

6300334

Interim Report

RESEARCH INFORMATION CENTER
NATIONAL BUREAU OF STANDARDS

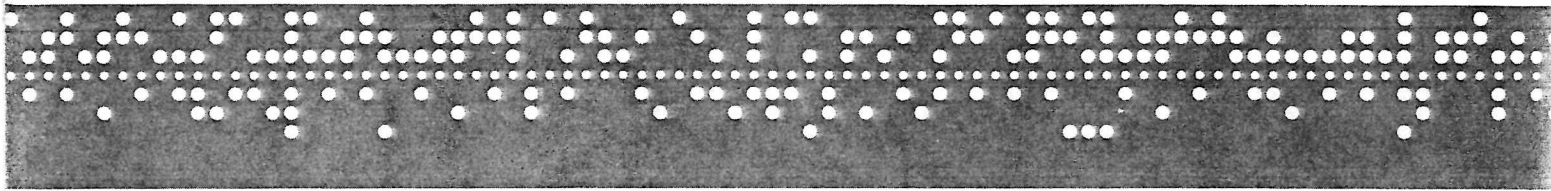
**MACHINE RECORDING of
TEXTUAL INFORMATION DURING
SCIENTIFIC JOURNAL PUBLICATION**

30 September 1963

Infor/onics inc.*

146 MAIN STREET • P.O. BOX 207 • MAYNARD, MASSACHUSETTS

10 NBS/217/3



The work described in this report was supported by the National Science Foundation under NSF Contract C-305.

ACCOMPLISHMENTS

Inforonics, Inc., is currently conducting a research program to develop publishing and computer processing techniques for recording useful textual data in machine form at the time of primary journal publication, so that it can be used for subsequent publishing and retrieval purposes.

Meeting this objective has required two developments:

■ ■ *A system for recording journal articles in a machine-interpretable form, so that the separate requirements of typographical composition, selective data extraction, and data retrieval are satisfied simultaneously by one keying. ■ ■ Transformation procedures to convert the recorded data to a form useful for information retrieval and secondary publication purposes.*

The purpose of this progress report is to present accomplishments in brief form. Further details of the project work are described in the preprints to the 1963 American Documentation Institute annual meeting, and a summary report is in publication.

SYSTEM DESCRIPTION

The system for recording and processing journal text data begins at the final stages of manuscript editing, when the manuscript is typed on a perforated-tape typewriter (illustrated on rear cover). After correction and proofreading, the tape is converted by a computer process to form both a typesetting tape for the journal article and published indexes, and a digital storage for subsequent uses. The typesetting tape is entered into a phototypesetter to produce typeset copy for making printing plates.

The following pages are samples of the system inputs and outputs, and were produced by the experimental system now in operation at Inforonics. The experimental results have demonstrated that:

- A single input keying of manuscripts in machine form can satisfy primary publication typesetting and also create, as a by-product, a machine record useful for information retrieval purposes, such as the compilation of indexes, abstract journals, and search files.
- A major portion of typographic layout of journal-type documents is machine-derivable from the identification of information items in the machine manuscript.
- All encoding of text required for either the identification of information items, the control of typographic form, or the selection of special symbols, can be accomplished with an ordinary keyboard, such as is available on standard perforated-tape typewriters.
- All input keying can be made independent of the typesetting machine being used, resulting in fewer training problems for typists. The input tapes prepared in this experiment can be used with any typesetting machine.
- The use of a typing format developed on this project is more economical than ordinary typesetting because it does not require extensive typographic control operations.

- The storage and use of a machine record of the input offers significant cost savings when the information is to be used repeatedly. The journal manuscript data is repeated in title pages and abstract journals during publication; and in the subsequent compilation and updating of author, subject and title indexes, the information is used even more repeatedly.

USE OF SYSTEM

A potential user of the system must analyze his publishing and information retrieval requirements to determine the following system specifications:

- A list of the types of text items which must be identified in the input record for use in primary or secondary publications.
- The output format of the printed publications.
- Range of type fonts which are required.
- A description of the text processing operations which must be performed prior to automatic typesetting, such as item extraction, conversion, sorting and merging.

Once these specifications have been developed, appropriate subroutines of the text processing program are selected and their format control and font tables modified to suit the requirements. A short sample publication containing examples of the full range of requirements is selected as a text sample, and is processed to uncover any errors before lengthy production runs are made.

FUTURE CAPABILITY

The ideas and concepts developed thus far will be extended to the production of other reference tools, such as subject and permuted title indexes and abstract journals. Also, the text information stored can be searched and processed for extraction or addition of data required for updating a standard reference tool or compilation. This latter capability will have broad use in the general problem of periodic publishing of compilations, such as catalogs and directories.

Spoken Digit Recognition Using Vowel-Consonant Segmentation

9.3 Acoustic Analysis of Speech 9.10 Machine Recognition of Speech

P. N. Sholtz and R. Bakis

IBM Research Center, Yorktown Heights, N. Y.

(Received September 11, 1961)

A procedure has been developed for recognition of spoken digits by means of digital computer simulation. Using power spectra computed at 10-msec intervals, the words are segmented into vowels and consonants. Vowels are then classified into one of 11 categories by a multivariate statistical decision method operating on approximations of the measurements. Consonants are classified into one of three categories by means of an empirically derived decision tree. Recognition is then performed by means of a dictionary search. When tested on a sample of 493 words spoken by 50 speakers, and with the internal dictionary adjusted for optimum results, 97% of the words were identified correctly. It appears that this procedure is more tolerant of interspeaker variations than those previously reported.

INTRODUCTION

This paper describes a procedure for automatic recognition of spoken digits. In recent years, several schemes for digit recognition have been described. Most of these have used either some type of matching technique or decision trees based upon empirically derived rules. The procedure to be described here is a combination of the empirical decision tree and a multivariate statistical decision method.

When tested on a sample of 483 words spoken by 50 speakers, and with the internal dictionary adjusted for optimum results with this sample, 96% of the words were identified correctly. All of the experimental work involved in the design and testing of this procedure was carried out by digital computer simulation.

DATA COLLECTION AND PREPARATION

The samples used throughout this experiment were uttered by 50 speakers, comprising 25 males and 25 females. The majority of the females spoke dialects typical of the New York City vicinity. For the males, the dialects were mostly of the varieties found in the northeastern and midwestern sections of the

Example of manuscript typed on a perforated-tape typewriter. Text items are identified by their sequence and position on page.

Spoken Digit Recognition Using Vowel-Consonant Segmentation

P. N. Sholtz and R. Bakis

IBM Research Center, Yorktown Heights, N. Y.

(Received September 11, 1961)

A procedure has been developed for recognition of spoken digits by means of digital computer simulation. Using power spectra computed at 10-msec intervals, the words are segmented into vowels and consonants. Vowels are then classified into one of 11 categories by a multivariate statistical decision method operating on approximations of the measurements. Consonants are classified into one of three categories by means of an empirically derived decision tree. Recognition is then performed by means of a dictionary search. When tested on a sample of 493 words spoken by 50 speakers, and with the internal dictionary adjusted for optimum results, 97% of the words were identified correctly. It appears that this procedure is more tolerant of interspeaker variations than those previously reported.

INTRODUCTION

This paper describes a procedure for automatic recognition of spoken digits. In recent years, several schemes for digit recognition have been described.^{1,2} Most of these have used either some type of matching technique or decision trees based upon empirically derived rules. The procedure to be described here is a combination of the empirical decision tree and a multivariate statistical decision method.

When tested on a sample of 483 words spoken by 50 speakers, and with the internal dictionary adjusted for optimum results with this sample, 96% of the words were identified correctly. All of the experimental work involved in the design and testing of this procedure was carried out by digital computer simulation.

DATA COLLECTION AND PREPARATION

The samples used throughout this experiment were uttered by 50 speakers, comprising 25 males and 25 females. The majority of the females spoke dialects typical of the New York City vicinity. For the males, the dialects were mostly of the varieties found in the northeastern and midwestern sections of the United States, although for two speakers English was not the native language. Speakers were instructed to speak the words carefully and naturally, but were given no training. They were merely instructed to pause between words.

All samples were recorded on magnetic tape, using a General Radio type 1551-P1 condenser microphone system, and an Ampex model 350-2 tape recorder. During the recording sessions, speakers were located in an acoustically insulated booth.

After recording, the speech signals were manually edited and digitalized for computer input. Two machines, the Editor and the Coder,⁴ were used for this purpose. By means of the Editor, pulses were recorded on a second track of the magnetic tape opposite the desired speech events. The tape was then played back by the Coder. There the signal was passed through an equalizer with a

Example of journal text galley prepared on Photon phototypesetter from input shown in previous example. Some characters in example are substituted because they were not available in any Photon type fonts.

Inner Ear Response to High-Level Sounds

Merle Lawrence, David Wolsk, and Pieter Schmidt*

Department of Otorhinolaryngology, The University of Michigan, Ann Arbor, Michigan

(Received June 22, 1961)

The cochlear ac potentials in response to a stimulating tone of rapidly increasing intensity undergo a rapid reduction in amplitude after reaching a certain maximum. The record seen on the cathode ray screen is indistinguishable from that reported for middle ear muscle action, yet the response described here occurs in

Studies of Nasal Consonants with an Articulatory Speech Synthesizer*

Michael H. L. Hecker

Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, Massachusetts

(Received September 19, 1961)

On the Width of Critical Bands

John A. Swets and David M. Green

Psychology Section and Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, Massachusetts

Wilson P. Tanner, Jr.

Cooley Electronics Laboratories, University of Michigan, Ann Arbor, Michigan

(Received August 24, 1961)

A different technique of analysis is applied to the experiment suggested by Harvey Fletcher for measuring the width of the critical band. This experiment determines the ability of noise bands of different widths to mask a pure tone centered in the band. The analysis considers two filters in series, one outside and one inside the observer. The width of the second filter (the critical band) can be estimated from measurements of the reduction in the noise power at the detector which is effected by the pair of filters. The width of the critical band is estimated under four different assumptions about the shape of the band. The results provide a context for discussing the reasons that may underlie the widely varying estimates of the critical bandwidth which have been obtained in previous studies.

INTRODUCTION

For the better part of a century, attempts to specify the process of auditory frequency analysis were based almost exclusively on anatomical and physiological evidence. Then, in 1940, Fletcher presented psychophysical data that gave a new form to the problem. He reported an experiment showing that only noise components in a narrow region about a pure tone are effective in masking the tone. This region he termed the "critical band."¹

an assumption since it was based on very few data, suggested that the critical band could be measured indirectly in masking experiments that used only broad-band noise. Fletcher later reported measurements based on broad-band noise; the critical bands so determined showed a similar dependence upon frequency and, again, the critical band in the region of 1000 cps was estimated to be approximately 65 cps wide.²

On the Width of Critical Bands

John A. Swets and David M. Green

Psychology Section and Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, Massachusetts

Wilson P. Tanner, Jr.

Cooley Electronics Laboratories, University of Michigan, Ann Arbor, Michigan

(Received August 24, 1961)

A different technique of analysis is applied to the experiment suggested by Harvey Fletcher for measuring the width of the critical band. This experiment determines the ability of noise bands of different widths to mask a pure tone centered in the band. The analysis considers two filters in series, one outside and one inside

Example of a different output journal format automatically typeset from same input tape as used on previous example.

Bakis, R. (see Sholtz, P. N.) 34; p.—1962.

Green, David M. (see Swets, John A.) 34; p.—1962.

Hecker, Michael H. L.. Studies of Nasal Consonants with an Articulatory Speech Synthesizer. 34; p.—1962.

Lawrence, Merle, David Wolsk, and Pieter Schmidt. Inner Ear Response to High-Level Sounds. 34; p.—1962.

Schmidt, Pieter. (see Lawrence, Merle) 34; p.—1962.

Sholtz, P. N., and R. Bakis. Spoken Digit Recognition Using Vowel-Consonant Segmentation. 34; p.—1962.

Swets, John A., and David M. Green. On the Width of Critical Bands. 34; p.—1962.

Wolsk, David. (see Lawrence, Merle) 34; p.—1962.

Example of author index entries produced automatically from the data tapes used in the preparation of the journal text samples. The page numbers are replaced with dashes; however, they will be included when the present typesetting program contains a page-numbering capability.

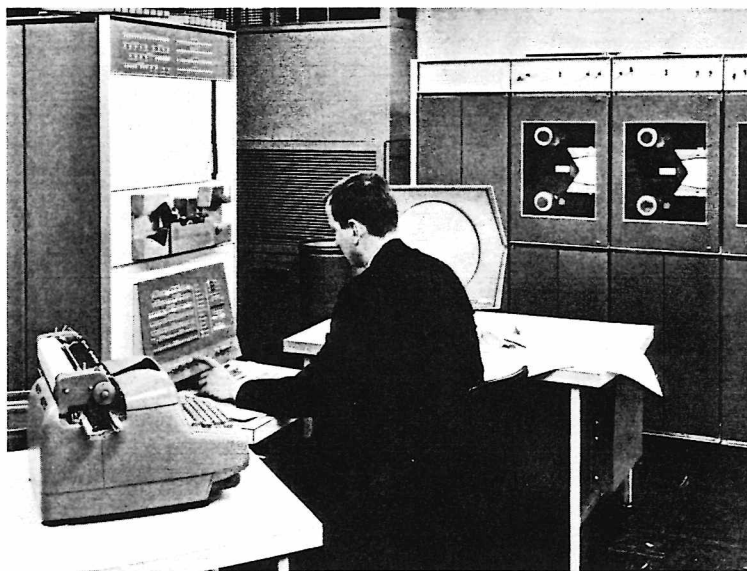
MANUSCRIPT TYPING

An operator prepares a manuscript on an ordinary perforated-tape typewriter. The copy is proofread and its associated tape is corrected prior to computer processing.



COMPUTER PROCESSING

A Digital Equipment Corporation PDP-1 computer is used to process the input tapes to produce journal typesetting tapes, index typesetting tapes, and a searchable data file.



OUTPUT PHOTOTYPESETTING

The tapes produced by the computer are entered into the Photon phototypesetter, pictured here, to produce galleys for the journal article and indexes. The samples in this report were produced at Machine Composition Company.

